Stability and Bifurcation for a Multiple Group Model for the Dynamics of HIV/AIDS Transmission

by

Wenzhang Huang,
Kenneth L. Cooke
and
Carlos Castillo-Chavez

BU-1071-M April 1990

Stability and Bifurcation for a Multiple Group Model for the Dynamics of HIV/AIDS Transmission

(Running Title: Stability and Bifurcation for an Epidemic Model)

Wenzhang Huang[1,2]

Center for Dynamical Systems and Nonlinear Studies
Georgia Institute of Technology, Atlanta, GA 30332

Kenneth L. Cooke[2]
Department of Mathematics
Pomona College, Claremont, CA 91711

Carlos Castillo-Chavez[2,3]
Biometrics Unit, Center for Applied Mathematics
341 Warren Hall, Cornell University, Ithaca, NY 14853-7801

Abstract. This paper examines a multi-group epidemic model with variable population size. We show that even in the case of proportionate mixing, multiple endemic equilibria are possible. The significance of these results in the study of the dynamics of sexually-transmitted diseases is discussed.

Key words. Acquired immunodeficiency syndrome (AIDS), Human immunodeficiency virus (HIV), Sexually transmitted diseases, proportionate mixing

AMS(MOS) subject classifications. 92A07, 92A15

---

[1]Most of the research reported herein constitutes part of this author's Ph.D. dissertation at the Claremont Graduate School.

[2]Partially supported by NSF Grant DMS 8807478.

[3]Partially supported by NSF Grant DMS 8906580 and NIAID Grant RO1 A129178-01. Correspondance or requests for reprints should be addressed to this author.

1. Introduction. The epidemic of AIDS, acquired immunodeficiency syndrome, is now in its second decade in the U. S., and over 124,000 cases have already been reported to the Centers for Disease Control (Atlanta, Georgia). The number of cases has continued to increase in Europe, Asia, and Latin America. The situation in some areas of Africa is devastating. Despite the severity of the situation there is not enough knowledge to forecast accurately the eventual magnitude of these epidemics. The nature of human interactions, the uncertainties in current estimates of epidemiological parameters, and the lack of enough reliable data, make it extremely difficult to understand the dynamics of AIDS without the frameworks provided by mathematical models. Furthermore, this increased knowledge is critically important in the development and evaluation of the relative merits of disease management programs (for an application of models to gonorrhea control, see [1]).

It has become increasingly clear that transmission of HIV, the causative virus for AIDS, depends on the social/sexual mixing structure of the population (with emphasis on transmission within and among core subgroups), on the long and variable incubation period (time between the point of acquisition of infection and the appearance of symptoms), and on variation of infectivity of an individual with the length of time since the individual was infected. A detailed discussion of the factors involved in the transmission of HIV can be found in [2-5], reviews of recent literature on modeling in [5-7], a comprehensive view of current issues related to AIDS epidemiology in [8].

In our previous work ([20-21]), we studied a single group model for the sexual transmission of HIV where the mixing is homogeneous. In Section 2 of this paper, we briefly recapitulate our analytical results on this model as we will be referring to them in our discussion of our results on multiple group models. The formulation of our multiple group models has been reported in [29-30]. The analysis presented in this paper is new.

The models analyzed here are for sexual transmission of HIV in a heterogeneously mixing population consist of an arbitrary number of distinct subgroups. The article emphasize the role in disease dynamics of the mixing matrix $\{p_{ij}\}$, where p_{ij} represents the fraction of partnerships of

individuals of group i that are with individuals of group j. The theoretical work in [13–16] and [38-39] shows that knowledge of the time-dependent mixing matrix can be of critical importance in disease dynamics. We show that models with variable population size can exhibit multiple endemic equilibria even in the case of random (proportionate) mixing. This result is somewhat surprising, since formally we are only dealing with a multigroup SIR (susceptible-infected-removed) model with variable population size. Models of this type have been simulated and partially analyzed in [3] and [9-10]. These simulations suggest the existence of a unique globally-stable equilibrium, but our analysis shows that this is not always the case. Although global stability may be true for the region of parameter space that is pertinent in AIDS transmission, we find either conclusion premature. In [17-18], the belief in the existence of a unique globally stable endemic equilibrium is also questioned, but on different grounds. We finally remark that our analysis does not contradict the analytical work of others. For example, if we remove most of the heterogeneties in our model, i.e. if we take most parameters to be equal (as it was done in [10]), then a tedious computation shows that you cannot have multiple endemic equilibria.

This paper is organized as follows: Section 2 describes the previous results for a single homogeneously mixing group model; Section 3 introduces our multiple group model for general and specific mixing functions; Section 4 discusses the stability of the disease-free equilibrium; Section 5 studies the existence of an endemic equilibrium, while its stability is discussed in Section 6. Section 7 proceeds with a global bifurcation analysis that shows, under appropriate conditions, the existence of multiple endemic equilibria. In Section 8, we discuss the relevance of these results and indicate future directions of research.

2. A basic model for a homogeneously mixing population. In this section, we review the formulation of a model for a homosexual population, with a single level of epidemiological heterogeneity, that introduced by a variable period of infectiousness. This model serves as a building block for the multiple-group models that are the main subject of this paper. The homosexually active population under consideration is divided into three groups: (S) (uninfected), I (infected by HIV,

asymptomatic or with mild symptoms), and A (HIV-infected, with severe symptoms). We make the simplifying assumption that all infected individuals are infectious (but see [17-18]) and that A-individuals (that is, individuals with "full-blown" AIDS) are sexually inactive. This last assumption is not essential, but is retained to simplify the discussion.

To describe the dynamics of HIV in this population, we let Λ denote the "recruitment" rate into the susceptible population, μ denote the rate of departure from sexual activity due to all causes except HIV infection, and λ denote the transmission rate per infectious partner. The parameter λ may be thought of as given by the product of two (uaually assumed) constant parameters, the average proportion of contacts with an infectious individual necessary for transmission, and the average number of contacts per sexual partner. C(T) denotes the mean number of sexual partners an average individual has per unit time, given that T is the total number of sexually-active persons ($T = S + I$). It is reasonable to expect that C(T) increases linearly when T is small and reaches a saturation level when T is large. Observing that I/T is the proportion of all partnerships that are formed with infected individuals, we arrive at the following expression for the incidence rate (new infections per unit time) B(t):

$$B(t) = \lambda C(T)S(T)\frac{I(t)}{T(t)}. \tag{1.1}$$

A variable period of infectiousness is modeled by the survivorship function P(s), which represents the proportion of those individuals who become infectious at time t and that, if not removed (that is, they survive as infectious), are still infectious at time $t + s$. P(s) is nonnegative,and nonincreasing, and $P(0) = 1$. We assume that

$$\int_0^\infty P(s)ds < \infty,$$

and observe that $-\dot{P}(s)$ gives the rate of removal from group I into group A, s time units after infection. Using these notations and definitions, we obtain the following distributed delay model for the sexual spread of HIV/AIDS (see [20-23]).

$$\frac{dS(t)}{d} = A - B(t) - \mu S(t)$$

$$I(t) = I_0(t) + \int_0^t B(x)e^{-\mu(t-x)}P(t-x)dx$$

$$A(t) = A_0(t) + A_1e^{-dt} + \int_0^t \left\{ \int_0^\sigma B(x)e^{-\mu(\sigma-x)}\left[-\dot{P}(\sigma - x)d^{-d(t-\sigma)}\right]dx \right\} d\sigma .$$

The functions $I_0(t)$, $A_0(t)$, and the constant A_1 are introduced to take into account initial conditions. The parameter d denotes the rate of removal of individuals from class A because of death.

In [20–21], we have shown that this model has at most two equilibria, which correspond to the infection-free state and the endemic state, and we have studied the stability of these equilibria. The parameter critical to the analysis of this model is the reproductive number, which may be defined as the number of secondary infections generated by a typical infectious individual in a population of susceptibles (for a rigorous definition of "typical" see [23]). For the model of this section, the reproductive number (a dimensionless parameter) is given by

$$R(T) = \lambda C(T)D ,$$

where

$$D = \int_0^\infty e^{-\mu s}P(s)ds$$

is the mean infectious period of sexually-active individuals. Using this dimensionless parameter, the following results were shown in [21]. First, the disease-free equilibrium is globally asymptotically stable, if and only if R_0 (the *basic* reproductive number, i.e. everybody is susceptible) $= R(\Lambda/\mu) \leq 1$. If $R_0 > 1$, then there is a unique endemic equilibrium, which is locally asymptotically stable. Further, when $P(s) = \exp(-\alpha s)$, the endemic equilibrium is globally asymptotically stable when it exists.

This model assumes that all infected individuals are equally infectious. However, recent studies ([24–26]) report that there is a high degree of variability in the level of virus titer present in HIV-infected persons, and it is thought that this correlates with variability in infectivity (a function of the age of the infection). Recent theoretical work ([17-18]) shows that under appropriate conditions, this

variability in infectivity is capable of forcing sustained oscillations in the total number of infecteds as well as in the incidence rate. Whether these oscillations can take place for parameter values that are epidemiologically relevant has yet to be investigated.

3. Formulation of n-group model. Although AIDS is not acquired at a constant rate, i.e., $P(s) = e^{-\alpha s}$ is an unrealistic survivorship function, we concentrate in the description (and later the analysis) of a HIV-transmission model with constant removal rates. Reasons for this decision include the fact that we are interested in studying the effects of different mixing patterns and hence begin with a model with few other complications, and also because this assumption does not limit the applicability of this model. It has been shown in [28–29] that realistic distributed delay can be incorporated by further subdividing the HIV-infected class and assuming constant transition rates between compartments. We feel that the incorporation of *this* type of distributed delay (generalized gamma distributions) will have no effect on the conclusions of this paper. However, we have not formally checked the validity of this claim. We further note that at present there is no analyses or clues as to the possible effects of varaible infectivity in multigroup models of this type.

Our basic multigroup model considers n sexually-active subpopulations each divided into the three epidemiological classes of Section 2: S_i, I_i, and A_i for $i = 1,2,\cdots,n$. Model parameters are defined as follows: Λ_i denotes the "recruitment" rate into S_i, μ denotes the removal rate from sexual activity due to all causes (except progression to AIDS), λ_{ij} denotes the transmission coefficients for passing the virus from a group j infective to a group i susceptible in one contact, and m_j denotes the proportion of contacts with a group j infectious individual needed for transmission. Further if W denotes a measure of the availability of sexual partners, which may be assumed to depend on $T_1,\cdots,T_n$ (where $T_j = S_j + I_j$), the total number of sexually-active individuals in each group then $c_i[W(T_1,\cdots T_n)]$ denotes the mean number of partners per unit time of a group i individual. To complete the model, we let $p_{ij}(t)$ denote the fraction of partners of a group i individual that are with group j individuals. Under these assumptions, the rate of infection of susceptibles in group i is given

by

$$B_i(t) = S_i(t)c_i[W(T_1,\cdots,T_n)] \sum_{j=1}^{n} \lambda_{ij} p_{ij}(t) \frac{I_j(t)}{T_j(t)} ,$$

and the equations of the model take the form

$$\frac{dS_i(t)}{dt} = \Lambda_i - B_i(t) - \mu S_i(t),$$

$$\frac{dI_i(t)}{dt} = B_i(t) - \mu(\sigma_i + 1)I_i(t),$$

$$\frac{dA_i(t)}{dt} = \alpha_i I_i(t) - (d_i + \mu)A_i(t),$$

where $\sigma_i = \alpha_i/\mu$, $i = 1,2,\cdots,n$.

The time-dependent mixing elements of the matrix $p_{ij}(t)$ are not arbitrary, since they must satisfy the following axioms for all times $t \geq 0$:

$$p_{ij} \geq 0, \qquad (i,j = 1,\cdots,n) ,$$

$$\sum_{j=1}^{n} p_{ij} = 1, \qquad (i = 1,\cdots,n) ,$$

$$c_i[W(T_1,\cdots,T_n)]T_i p_{ij} = c_j[W(T_1,\cdots,T_n)]T_j p_{ji}, \qquad (i,j = 1 = 1,\cdots,n) .$$

Although a general representation formula has been found recently for all solutions of the above axioms (see [15-16]), in this paper we shall consider a special case, the so-called preferred mixing (see [9-10], [29–30], [39]), in which

$$p_{ij} = \begin{cases} \eta_i + (1 - \eta_i)c_i[W(T_1,\cdots,T_n)](1 - \eta_i)T_i/N(T_1,\cdots,T_n), & i = j \\ (1 - \eta_i)c_j[W(T_1,\cdots,T_n)](1 - \eta_j)T_j/N(T_1,\cdots,T_n), & i \neq j \end{cases}$$

where $i,j = 1,\cdots,n$ and

$$N(T_1,\cdots,T_n) = \sum_{k=1}^{n} (1 - \eta_k)T_k c_k[W(T_1,\cdots,T_n)] .$$

The dependence of the T_j and p_{ij} on t has not been indicated in these formulas for notational simplicity. In this definition, η_i denotes the fraction of group i's contacts that are reserved for within the i'th subpopulation, while the remaining fraction, $1 - \eta_i$, are assumed to be distributed according to proportional mixing within all n groups. These fractions are assumed to be constants. In case all

$\eta_i = 0$, this reduces to the so-called proportionate or random mixing case. Despite the fact that selecting time-independent η_i's imposes specific time-dependent behavioral changes (see [39]), we still use preferred mixing because it has been used extensively, with some degree of success, in the mathematical epidemiological literature ([8],[10-14],[20]). We note however, that the results of this paper hold even for the most basic form of mixing, that is, proportionate or random mixing.

It is now convenient to transform the equations of the model by substituting the expressions for p_{ij} into the previous equations. We shall restrict attention in this analysis to the case in which the c_i are constants, independent of the size of the various populations. This is a reasonable approximation until such time as there is a substantial change in sexual behavior (see [29]). If we define

$$\theta_i = \eta_i \lambda_{ii} c_i, \qquad r_i = c_i(1 - \eta_i),$$

$$l_{ij} = c_i(1 - \eta_i)c_j(1 - \eta_j)\lambda_{ij} = r_i r_j \lambda_{ij},$$

$$N(T) = \sum_{k=1}^{n} c_k T_k (1 - \eta_k) = \sum_{k=1}^{n} r_k T_k ,$$

the resulting equations take the form

$$\frac{dS_i}{dt} = \Lambda_i - S_i\left(\frac{\theta_i I_i}{T_i} + \frac{1}{N(T)} \sum_{j=1}^{n} l_{ij} I_j \right) - \mu S_i = X_i(S, I),$$

$$\frac{dI_i}{dt} = S_i\left(\frac{\theta_i I_i}{T_i} + \frac{1}{N(T)} \sum_{j=1}^{n} l_{ij} I_j \right) - \mu(\sigma_i + 1) I_i = Y_i(S, I), \qquad (3.1)_\mu$$

$$i = 1,\cdots,n$$

where $S = (S_1,\cdots,S_n)^T \in R^n$, $I = (I_1,\cdots,I_n)^T \in R^n$, $\mu > 0$, $\theta_i \geq 0$, $l_{ij} \geq 0$, $\Lambda_i > 0$, $\sigma_i > 0$ are constants, and $N(T) = \sum_{k=1}^{n} r_k T_k$. We no longer consider the equations for A_i since they play no role in the mathematical analysis.

In the rest of this paper we investigate the existence and stability of nonnegative equilibria (from a biological point of view only nonnegative solutions make sense) and the bifurcation problem for the system $(3.1)_\mu$ with μ considered as the parameter.

4. The stability of the disease free equilibrium. To analyze $(3.1)_\mu$ it is sometimes convenient to let

$$X(\mu, S, I) = (X_1,\cdots,X_n)^T, \quad Y(\mu, S, I) = (Y_1,\cdots,Y_n)^T ,$$

and write $(3.1)_\mu$ in the vector form

$$\frac{dS}{dt} = X(\mu, S, I), \quad \frac{dI}{dt} = Y(\mu, S, I) . \qquad (4.1)_\mu$$

The system $(3.1)_\mu$ always has a trivial equilibrium:- $-(S_0, I_0) = (\Lambda_1/\mu,\cdots,\Lambda_n/\mu,0,\cdots,0)-$ –which is called the disease-free state.

The Jacobi matrix at (S_0, I_0) is given by the expression

$$\frac{\partial}{\partial(S, I)}\begin{bmatrix} X(\mu, S_0, I_0) \\ Y(\mu, S_0, I_0) \end{bmatrix} = \begin{bmatrix} -\mu E & -H \\ 0 & H - \mathrm{diag}[\mu(\sigma_i + 1)] \end{bmatrix},$$

where $E = E_{n\times n}$ is the identity matrix , H is the $n\times n$ matrix

$$H = \mathrm{diag}(\theta_i) + \frac{1}{K}\mathrm{diag}(\Lambda_i)L, \qquad K = \sum r_k\Lambda_k ,$$

and $L = [l_{ij}]_{n\times n}$

Remark: Throughout this paper we assume that L is irreducible (see [34] for the definition of "irreducibility").

The main stability result of this paper is the following:

Theorem 4.1: *Let $\rho(A)$ denote the spectral radius of $A \in R^{n\times n}$ and let*

$$\mu_0 = \rho\left[diag\left(\frac{\theta_i}{(\sigma_i + 1)}\right) + diag\left(\frac{\Lambda_i}{K(\sigma_i + 1)}\right)L \right],$$

then the disease free state is locally asymptotically stable if $\mu > \mu_0$ and unstable if $\mu < \mu_0$.

Proof: It is clear that (S_0, I_0) is locally asymptotically stable if all eigenvalues of $H - \mathrm{diag}[\mu(\sigma_i + 1)]$ have negative real parts, and unstable if $H - \mathrm{diag}[\mu(\sigma_i + 1)]$ has an eigenvalue with positive real part. Since H is a nonnegative matrix, our assertion follows from the well-known result in the theory of nonnegative matrices ([33, p. 12]).

5. **Local bifurcation and the existence of the positive endemic equilibrium.** It is natural to ask whether equation $(3.1)_\mu$ has a positive equilibrium. To address this question, we observe that if [S(t), I(t)] is a solution of $(3.1)_\mu$ with initial condition $S(0) \geq 0$, $I(0) \geq 0$, then $S(t) \geq 0$, $I(t) \geq 0$ for all $t \geq 0$ and

$$\limsup_{t\to\infty}[S_i(t) + I_i(t)] \leq \frac{\Lambda_i}{\mu}. \qquad i = 1,\cdots,n$$

Hence $(3.1)_\mu$ is positively invariant and dissipative. Since the trival solution $(\Lambda/\mu,\ 0)$ $[\Lambda=(\Lambda_i,\cdots,\Lambda_n)^T]$ is located at the boundary of R^{2n}_+, and since the stability of this trival equilibrium switches as the parameter μ passes through μ_0, then it is reasonable to expect the occurrence of a positive equilibrium for values of μ near μ_0

This section addresses the bifurcation at $\mu = \mu_0$ associated with the existence of positive equilibria by looking for positive solutions of the nonlinear system

$$X(\mu, S, I) = 0, \qquad Y(\mu, S, I) = 0\ . \tag{5.1}_\mu$$

By adding X to Y it is easy to see that the problem of the existence of such positive solutions is equivalent to the solution of the following system:

$$Y\left[\mu, \frac{\Lambda}{\mu} - (\sigma + E)I, I\right] = 0, \qquad S = \frac{\Lambda}{\mu} - (\sigma + E)I \gg 0,\ I \gg 0\ .$$

where the notation "$x \gg y$" means that each component of x is strictly larger than the corresponding component of y, $\sigma = \mathrm{diag}(\sigma_i)$ and E is the identity matrix. Explicitly, this is equivalent to the problem of finding solutions I of

$$F(\mu, I) \overset{def}{=} \left[\sum_{k=1}^{n} r_k\left(\frac{\Lambda_k}{\mu} - \sigma_k I_k\right)\right]\mathrm{diag}\left(\frac{1}{\Lambda_i/\mu - (\sigma_i + 1)I_i}\right)Y\left(\mu, \frac{\Lambda}{\mu} - (\sigma + E)I, I\right) = 0 \tag{5.2}_\mu$$

with

$$I \in B(\mu) \overset{def}{=} \{I = (I_1, \cdot, I_n)^T \in R^n : 0 < I_i < \Lambda_i/\mu(\sigma_i + 1), \quad i = 1,\cdots,n\}\ ,$$

and then setting $S = \Lambda/\mu - (\sigma + E)I$.

A further calculation shows that

$$F(\mu, I) = G(\mu, I)I$$

with

$$G(\mu,I) = L - g(\mu,I)\ \mathrm{diag}(y_i(\mu, I_i)),$$

where

$$g(\mu,I) = K - \mu \sum_{k=1}^{n} r_k \sigma_k I_k, \qquad K = \sum r_k \Lambda_k$$

and

$$y_i(\mu, I_i) = \frac{\mu(\sigma_i + 1)}{\Lambda_i - \mu(\sigma_i + 1)I_i} - \frac{\theta_i}{\Lambda_i - \mu\sigma_i I_i}. \qquad i = 1,\cdots,n$$

In addition, since

$$\rho\left[\mathrm{diag}\left(\frac{\theta_i}{\sigma_i + 1}\right) + \mathrm{diag}\left(\frac{\Lambda_i}{K(\sigma_i + 1)}\right)L\right] = \mu_0 ,$$

there is an $I_0 = (I_{01},\cdots,I_{0n})^T \in R_+^n$ such that

$$\left[\mathrm{diag}\left(\frac{\theta_i}{\sigma_i + 1}\right) + \mathrm{diag}\left(\frac{\Lambda_i}{K(\sigma_i + 1)}\right)L\right] = \mu_0 I_0 .$$

It follows that

$$G(\mu_0,0)I_0 = 0 .$$

A straightforward application of M-matrix theory ([34, p. 156]) shows that $G(\mu_0,0)$ is a singular M-matrix. Since $G(\mu_0,0)$ is irreducible, and 0 is a simple eigenvalue of $G(\mu_0,0)$, it follows that

$$R^n = N[G(\mu_0,0)] \oplus \mathrm{Ran}[G(\mu_0,0)] ,$$

where N and Ran denote the null space and range of , respectively. Consequently, the mapping

$$G(\mu_0,0)_{|\mathrm{Ran}[G(\mu_0,0)} : \mathrm{Ran}[G(\mu_0,0)] \rightarrow \mathrm{Ran}[G(\mu_0,0)]$$

is one-to-one and onto.

If $0 \neq \alpha \in R$ and $z \in \mathrm{Ran}[G(\mu_0,0)]$ are such that

$$G[\mu,\alpha(I_0 + z)](I_0 + z) = 0 , \tag{5.3}$$

then

$$F[\mu,\alpha(I_0 + z)] = 0 \qquad \text{and} \qquad \alpha(I_0 + z) \neq 0 .$$

Since it is clear that $G^T(\mu_0,0)$ is also a singular M-matrix, we can find $I_0^* = (I_{01}^*,\cdots,I_{0n}^*)^T \in$

R_+^n such that

$$G^T(\mu_0,0)I_0^* = 0, \qquad \langle\, I_0^*,\, I_0 \,\rangle = 1\,,$$

where $\langle\, x,y \,\rangle = \sum x_j y_j$ for x, y $\in R^n$. As z $\in$ Ran[$G(\mu_0,0)$] if and only if $\langle\, I_0^*,\, z \,\rangle = 0$, then we can define the operator $T : R_+ \times R \times R^n \to R \times R^n$ by

$$T(\mu,\, \alpha,\, z) = \{\langle\, I_0^*,\, z \,\rangle,\, G[\mu,\, \alpha(I_0 + z)](I_0 + z)\}.$$

Equation (5.3) is now equivalent to

$$T(\mu,\, \alpha,\, z) = 0, \tag{5.4}$$

$$T(\mu_0,0,\, 0) = [0, G(\mu_0,0)I_0] = 0,$$

and

$$\frac{\partial T(\mu_0,0,\, 0)}{\partial(\alpha,\, z)}\,(\alpha,z) = [\langle\, I_0^*,\, z \,\rangle,\, G(_0,0)z - \alpha\, \mathrm{diag}(s_i)I_0],$$

where

$$x_i = \frac{d}{d\alpha}\,[g(\mu_0,\alpha I_0)y_i(\mu_0,\alpha I_{0i})]\big|_{\alpha=0}$$

$$= \mu_0\left[\frac{K\, I_{0i}[\mu_0(\sigma_i + 1)^2 - \theta_i\sigma_i)}{\Lambda_i^2} - \left(\sum_{j=1}^{n} r_j\sigma_j I_{0j}\right)\frac{\mu_0(\sigma_i + 1) - \theta_i}{\Lambda_i}\right].$$

Finally using the expression

$$h(\mu_0) = \langle\, I_0^*,\, \mathrm{diag}(x_i)I_0 \,\rangle = \sum_{j=1}^{n} x_j I_{0j}^* I_{0j}\,,$$

we arrive at the following result:

Lemma 5.1: *Suppose* $h(\mu_0) \neq 0$, *then*

$$\frac{\partial T(\mu_0,0,\, 0)}{\partial(\alpha,\, z)} : R \times R^n \to R \times R^n$$

is one-to-one and onto.

Proof: For arbitrary (b, y) $\in R \times R^n$, we let $\alpha = -\langle\, I_0^*,\, y \,\rangle / h(\mu_0)$. Hence

$$\langle\, I_0^*,\, \alpha\, \mathrm{diag}(x_i)I_0 + y \,\rangle = 0,$$

and $\alpha \mathrm{diag}(x_i) I_0 + y \in \mathrm{Ran}[G(\mu_0,0)]$. There is a unique $z_0 \in \mathrm{Ran}[G(\mu_0,0)]$ such that

$$G(\mu_0,0)z_0 = \alpha\, \mathrm{diag}(x_i) I_0 + y\ ,$$

or

$$G(\mu_0,0)z_0 - \alpha\, \mathrm{diag}(x_i) I_0 = y\ .$$

Since $z_0 \in \mathrm{Ran}[G(\mu_0,0)]$, we have $\langle\, I_0^*,\, z_0\,\rangle = 0$. Let $z = bI_0 + z_0$, note that $\langle\, I_0^*,\, I_0\,\rangle = 1$, then

$$\langle\, I_0^*,\, z\,\rangle = \langle\, I_0^*,\, bI_0 + z_0\,\rangle = b\ .$$

Therefore

$$\frac{\partial T(\mu_0,0,0)}{\partial(\alpha,\, z)}\,(\alpha,\, z) = (b,\, y)\ .$$

Moreover it is easy to see that if $(b,\, y) = (0,\, 0)$, then $(\alpha,\, z) = (0,\, 0)$, and hence $\partial T(\mu_0,0,\, 0)/\partial(\alpha,\, z)$ is one-to-one and onto.

With the aid of Lemma 5.1, we arrive at the following:

Theorem 5.2: *Suppose $h(\mu_0) \neq 0$, then μ_0 is a bifurcation point. Specifically, if $h(\mu_0) > 0$ (<0), then there exist $\epsilon > 0$ and a continuously differentiable function*

$$(S,\, I) : (\mu_0 - \epsilon,\, \mu_0]\ ([\mu_0,\, \mu_0 + \epsilon)) \rightarrow R^{2n}$$

such that $[S(\mu_0),\, I(\mu_0)] = (\Lambda/\mu_0,0)$, $[S(\mu),\, I(\mu)]$ is strictly positive for $\mu \neq \mu_0$, and

$$F[\mu,\, I(\mu)] = X[\mu,\, S(\mu),\, I(\mu)] = Y[\mu,\, S(\mu),\, I(\mu)] = 0, \qquad \mu \in (\mu - \epsilon,\, \mu_0]([\mu_0,\mu_0 + \epsilon))\ .$$

Proof: It follows from Lemma 5.1 and the implicit function theorem that there are $\epsilon > 0$ and a continuously differentiable function $(\alpha,z) : J = (\mu_0 - \epsilon,\, \mu_0 + \epsilon) \rightarrow R \times \mathrm{Ran}[G(\mu_0,0)]$ such that $[\alpha(\mu_0),\, z(\mu_0)] = (0,\, 0)$ and

$$G[\mu,\, \alpha(\mu)][I_0 + z(\mu)][I_0 + z(\mu)] = 0, \qquad \mu \in J\ . \tag{5.5}$$

After differentiating (5.5) with respect to μ at μ_0 and using the chain rule, one obtains

$$G(\mu_0,0)\,\frac{dz(\mu_0)}{d\mu} - \frac{d\alpha(\mu_0)}{d\mu}\,\mathrm{diag}(x_i) I_0 = -\frac{\partial G(\mu_0,0)}{\partial \mu} I_0 = \mathrm{diag}[K(\sigma_i + 1)/\Lambda_i] I_0\ . \tag{5.6}$$

Use of the relation

$$\langle\, I_0^*,\, G(\mu_0,0)\,\frac{dz(\mu_0)}{d\mu}\,\rangle = \langle\, G^T(\mu_0,0) I_0^*,\, \frac{dz(\mu_0)}{d\mu}\,\rangle = 0$$

(5.6) implies that

$$-\frac{d\alpha(\mu_0)}{d\mu}h(\mu_0) = -\frac{d\alpha(\mu_0)}{d\mu}\langle I_0^*, \mathrm{diag}(x_i)I_0\rangle = \langle I_0^*, \mathrm{diag}[K(\sigma_i + 1)/\Lambda_i]I_0\rangle \gg 0 .$$

Therefore, we have

$$\frac{d\alpha(\mu_0)}{d\mu}\begin{cases} > 0 & \text{if } h(\mu_0) < 0 \\ < 0 & \text{if } h(\mu_0) > 0 \end{cases} . \tag{5.7}$$

Note that $\alpha(\mu_0) = 0$; so if ϵ is small enough then

$$\alpha(\mu) > 0, \quad \mu \in (\mu_0 - \epsilon)[(\mu_0, \mu_0 + \epsilon)] \quad \text{if} \quad h(\mu_0) > 0[h(\mu_0) < 0] .$$

Since I_0 is strictly positive and $z(\mu_0) = 0$, if ϵ is sufficiently small, then

$$0 \ll \alpha(\mu)[I_0 + z(\mu)] \in B(\mu), \ \mu \in (\mu_0 - \epsilon, \mu_0)\ [(\mu_0, \mu_0 + \epsilon)] \quad \text{if} \quad h(\mu_0) > 0[h(\mu_0) < 0] .$$

Consequently, $I(\mu) = \alpha(\mu)[I_0 + z(\mu)]$ is a positive solution of $F(\mu, I) = 0$, and $[S(\mu), I(\mu)]$ is a continuously differentiable positive equilibrium of system $(3.1)_\mu$. Explicitly,

$$S_i(\mu) = \Lambda_i/\mu - (\sigma_i + 1)I_i(\mu). \quad i = 1,\cdots,n$$

6. Stability of the positive equilibrium. We now turn to the investigation of the stability of the positive equilibrium bifurcating from $\mu = \mu_0$. We begin by stating the following lemma:

Lemma 6.1: *Let $Z(\mu)$ be a $2n \times 2n$ real matrix satisfying the following conditions:*

(*i*) *$Z(\mu)$ is continuous for $\mu \in (\mu_0 - \epsilon, \mu_0 + \epsilon)$.*

(*ii*) *$Z(\mu_0)$ has a zero eigenvalue which is simple and all other eigenvalues of $Z(\mu_0)$ have negative real parts.*

If $det[Z(\mu)] > 0$, then hen all eigenvalues of $Z(\mu)$ have negative real parts , while if $detZ(\mu) < 0$, then $Z(\mu)$ has a positive eigenvalue, where $|\mu - \mu_0|$ is sufficiently small.

This lemma is adirect consequence of the continuous dependence of the eigenvalues of $Z(\mu)$ on μ. We omit the proof.

Lemma 6.2: *Let $I(\mu)$ be the positive solution of equation $(5.2)_\mu$ defined in Theorem* 5.2. *Then*

$$det\left(\frac{\partial F(\mu, I(\mu))}{\partial I}\right) > 0 \quad \text{if} \quad h(\mu_0) > 0, \quad \mu \in (\mu_0 - \epsilon, \mu_0) ,$$

and

$$det\left(\frac{\partial F(\mu,I(\mu))}{\partial I}\right) < 0 \quad if \quad h(\mu_0) < 0, \qquad \mu \in (\mu_0,\ \mu_0 + \epsilon)\ .$$

Proof: Differentiation of $F[\mu,\ I(\mu)] = 0$ implies that

$$\{-\ D_I F[\mu,\ I(\mu)]\}\left(-\frac{dI(\mu)}{d\mu}\right) = -D_\mu F[\mu,\ I(\mu)], \tag{6.1}$$

and from the definition of $F(\mu,\ I)$, one verifies

$$D_I[\mu,\ I(\mu)] = G(\mu,\ I) + \mu \mathrm{diag}[I_i y_i(\mu,\ I_i)]Q - g(\mu,\ I)\mathrm{diag}\left(\frac{\partial y_i(\mu,\ I_i)}{\partial I_i}\right).$$

Here Q is an n×n matrix with entries $Q_{ij} = r_j\sigma_j$, and $I = I(\mu)$. Hence it is easy to see that all the off-diagonal entries of $-D_I F(\mu,\ I)$ are nonpositive. Moreover we have

$$-\ D_\mu F_i[\mu,\ I(\mu)] = I_i(\mu)\left[\frac{\partial g(\mu,\ I)}{\partial \mu}\ y_i(\mu,\ I_i) + g(\mu,\ I)\ \frac{\partial y_i(\mu,\ I_i)}{\partial \mu}\right],$$

and

$$\frac{\partial g(\mu,\ I)}{\partial \mu}\ y_i(\mu,\ I_i) + g(\mu,\ I)\ \frac{\partial y_i(\mu,\ I_i)}{\partial \mu}_{|[\mu_0, I(\mu_0)]} = \frac{K(\sigma_i + 1)\Lambda_i}{\Lambda_i^2} > 0\ ,$$

where F_i is the *ith* component of F. The continuity of $I(\mu)$ implies

$$-\ D_\mu F_i[\mu,\ I(\mu)] > 0, \qquad i = 1,\cdots,n \tag{6.2}$$

if $h(\mu_0) > 0$ $[h(\mu_0) < 0]$ and $\mu \in (\mu_0 - \epsilon,\ \mu_0)[\mu \in (\mu_0,\ \mu_0 + \epsilon)]$.

Finally,

$$I(\mu) = \alpha(\mu)[I_0 + z(\mu)], \qquad \alpha(\mu_0) = 0,\ z(\mu_0) = 0,$$

and therefore, (5.7) yields

$$-\frac{dI(\mu_0)}{d\mu} = -\frac{d\alpha(\mu_0)}{d\mu}\ I_0 \gg 0\ (\ll 0) \quad \text{if } h(\mu_0) > 0\ [h(\mu_0) < 0]\ . \tag{6.3}$$

We look at the two cases:

(i) Suppose $h(\mu_0) > 0$. The continuity of $dI(\mu)/d\mu$, (6.2), and (6.3) imply that

$$-\ \frac{dI(\mu)}{d\mu} \gg 0, \qquad -\ D_\mu F[\mu,\ I(\mu)] \gg 0, \qquad \mu \in (\mu_0 - \epsilon,\ \mu_0)\ .$$

Hence, M-matrix theory ([34, p. 141]) and (6.1) imply that $-D_I F[\mu, I(\mu)]$ is a nonsingular M-matrix, and consequently, $\det\{-D_I F[\mu, I(\mu)]\} > 0$.

(ii) If $h(\mu_0) < 0$, then we have $-\, dI(\mu)/d\mu \ll 0$. First notice that

$$-\, D_I F[\mu_0, I(\mu_0)] = -\, G(\mu_0, 0)\ .$$

Since $-G(\mu_0, 0)$ is an irreducible singular M-matrix, all leading principal minors of $-D_I F[\mu, I(\mu)]$ other than $-D_I F[\mu, I(\mu)]$ are positive if $\mu \in (\mu_0, \mu_0 + \epsilon)$ and ϵ is small enough. Hence if $\det\{-D_I F[\mu, I(\mu)]\} \geq 0$, then $-\, D_I F[\mu, I(\mu)]$ would be an M-matrix; therefore, it follows from [34, p. 137] that

$$-\, \frac{dI(\mu)}{d\mu} \gg 0\ .$$

This is a contradiction and the proof of this lemma is completed.

We are now ready to prove the main result of this section.

Theorem 6.3: *The positive equilibrium* $[S(\mu), I(\mu)]$ *of* $(1.2)_\mu$ *is locally asymptotically stable (unstable) if* $h(\mu_0) > 0$ $[h(\mu_0) < 0]$ *and* $\mu \in (\mu_0 - \epsilon, \mu_0)$ $[\mu \in (\mu_0, \mu_0 + \epsilon)]$. *Here* ϵ *is a sufficiently small positive number.*

Proof: Let

$$Z(\mu, S, I) = \begin{bmatrix} X(\mu, S, I) \\ Y(\mu, S, I) \end{bmatrix}.$$

By the definition of X and Y, we have

$$X(\mu,S,I) + Y(\mu,S,I) = \Lambda - \mu S - \mu(\sigma + E)I,\ \ S(\mu) = \frac{\Lambda}{\mu} - (\sigma + E)I(\mu),\ \ Z[\mu,S(\mu),I(\mu)] = 0\ ,$$

and

$$F(\mu, I) = Y_1\left(\mu, \frac{\Lambda}{\mu} - (\sigma + E)I, I\right),$$

where

$$Y_1(\mu, S, I) = N(T)\mathrm{diag}\left(\frac{1}{S_i}\right)Y(\mu, S, I)\ .$$

Hence,

$$D_I F(\mu, I) = -\, D_S Y_1\left(\mu, \frac{\Lambda}{\mu} - (\sigma + E)I, I\right)[\sigma + E] + D_I Y_1\left(\mu, \frac{\lambda}{\mu} - (\sigma + E)I, I\right);$$

therefore,

$$\begin{bmatrix} E & E \\ 0 & N[T(\mu)]\mathrm{diag}[1/X_i(\mu)] \end{bmatrix} \frac{\partial Z[\mu,S(\mu),I(\mu)]}{\partial(I,S)} \begin{bmatrix} E & -(\sigma+E) \\ 0 & E \end{bmatrix}$$

$$= \frac{\partial}{\partial(S,I)} \left(\begin{bmatrix} E & E \\ 0 & N(T)\mathrm{diag}(1/S_i) \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \right) \begin{bmatrix} E & -(\sigma+E) \\ 0 & E \end{bmatrix} \Big|_{[\mu,S(\mu),I(\mu)]}$$

$$= \frac{\partial}{\partial(S,I)} \begin{bmatrix} \Lambda - \mu S - \mu(\sigma+E) \\ Y_1(\mu,S,I) \end{bmatrix} \Big|_{[\mu,S(\mu),I(\mu)]} \begin{bmatrix} E & -(\sigma+E) \\ 0 & E \end{bmatrix}$$

$$= \begin{bmatrix} -\mu E & 0 \\ -D_S Y_1 & -D_S Y_1(\sigma+E) + D_I Y_1 \end{bmatrix} \Big|_{[\mu,S(\mu),I(\mu)]}$$

$$= \begin{bmatrix} -\mu E & 0 \\ -D_S Y_1[\mu,S(\mu),I(\mu)] & D_I F[\mu,S(\mu),I(\mu)] \end{bmatrix}.$$

It now follows that

$$N[T(\mu)]^n \prod_{i=1}^{n} S_i^{-1} \det \frac{\partial Z[\mu,S(\mu),I(\mu)]}{\partial(S_i,I)} = \mu^n \det\{-D_I F[\mu,I(\mu)]\}, \tag{6.4}$$

where $N[T(\mu)] = \sum_{k=1}^{n} r_k[S_k(\mu) + I_k(\mu)]$. Fom (6.4) and Lemma 6.2 we see

$$\det\left[\frac{\partial Z[\mu,S(\mu),I(\mu)]}{\partial(S,I)}\right] \begin{cases} > 0 & \text{if } h(\mu_0) > 0 \text{ and } \mu \in (\mu_0 - \epsilon, \mu_0) \\ < 0 & \text{if } h(\mu_0) < 0 \text{ and } \mu \in (\mu_0, \mu_0 + \epsilon) \end{cases}.$$

Since all eigenvalues of $\partial Z[\mu_0, S(\mu_0), I(\mu_0)]/\partial(S, I)$ have negative real parts, except for a simple zero eigenvalue, then by applying Lemma 6.1 we conclude the proof of our theorem.

7. Global bifurcation. In general, it is difficult to know exactly how many positive equilibria the system $(3.1)_\mu$ has. But by using the information from the local bifurcation analysis in Section 4 we can show that $(3.1)_\mu$ has (under certain assumptions) at least two endemic (i. e. positive) equilibria. To establish the possibility of multiple endemic equilibria, we proceed to study the global bifurcation from $\mu = \mu_0$. We begin by introducing four technical lemmas that are used later on.

Lemma 7.1: *For $\mu > 0$, let $B(\mu)$ be defined as in Section 5. There is μ^*, sufficiently large, such that*

$$F(\mu, I) \neq 0, \qquad I \in B(\mu), \qquad \mu \geq \mu^* .$$

Proof: If $I \in B(\mu)$, then for each i,

$$g(\mu, I)y_i(\mu, I_i) \geq g(\mu, I)\frac{\mu(\sigma_i + 1) - \theta_i}{\Lambda_i - \mu(\sigma_i + 1)I_i} \geq \mu(\sigma_i + 1) - \theta_i,$$

hence,

$$g(\mu, I)y_i(\mu, I_i) \to \infty \quad \text{as} \quad \mu \to \infty .$$

Therefore, there is $\mu^* > 0$ such that

$$g(\mu, I)y_i(\mu, I_i) > \rho(L), \qquad \mu \geq \mu^*, \qquad I \in B(\mu), \qquad i = 1,\cdots,n$$

so

$$F(\mu, I) = [L - g(\mu, I)\text{diag}y_i(\mu, I_i)]I \neq 0. \qquad \mu \geq \mu^*, \qquad I \in B(\mu)$$

Lemma 7.2: *For each ϵ, there is $\delta(\epsilon)$ 0 such that*

$$F(\mu, I) \neq 0 \qquad \textit{for all} \qquad \mu \in J_\epsilon = [0, \mu_0 - \epsilon] \cup [\mu_0 + \epsilon, \mu^*], \qquad I \in B_\delta ,$$

where

$$B_\delta = \{I \in R^n, 0 < I_i < \delta, \qquad i = 1,\cdots,n\} .$$

Proof: Suppose the contrary; then there is a sequence $\{(\mu_n, I^n)\}$, such that $\mu_n \in J_\epsilon$, $\|I^n\| \to 0$ as $n \to \infty$, and $F(\mu_n, I^n) = 0$. Without loss of generality, assune that $\mu_n \to \mu^0 \in J_\epsilon$, and note that

$$\begin{aligned} 0 &= \langle I_0^*, F(\mu_n, I^n) \rangle \\ &= \langle I_0^*, G(\mu_n, I^n)I^n \rangle \end{aligned}$$

$$= \langle I_0^*, G(\mu_0, 0)I^n + [G(\mu_n, I^m) - G(\mu_0, 0)]I^n \rangle \tag{7.1}$$

$$= \langle I_0^*, [G(\mu_n, I^n) - G(\mu_0, 0)]I^n \rangle .$$

Further, $\|I^n\| \to 0$ implies that

$$\lim_{n\to\infty}[G(\mu_n, I^n) - G(\mu_0, 0)] = -K \operatorname{diag}[y_i(\mu^0, 0) - y_i(\mu_0, 0)]$$

$$= -K \operatorname{diag}\left(\frac{\sigma_i + 1}{\Lambda_i}(\mu^0 - \mu_0)\right). \tag{7.2}$$

Since $\mu^0 \neq \mu_0$ and $I_0^* \gg 0$, $I^n \gg 0$, $n = 1,\cdots$, (7.2) implies that for large enough n

$$\langle I_0^*[G(\mu_n, I^n) - G(\mu_0, 0)]I^n > \neq 0 .$$

This contradicts (7.1), completing the proof.

Now let

$$U(\mu, I) = \operatorname{diag}\left(\frac{\Lambda_i - (\sigma_i + 1)I_i}{\mu^2 g(\mu, I)(\sigma_i + 1)}\right) F(\mu, I) .$$

It is apparent that for $I \in B(\mu)$, $\mu > 0$

$$F(\mu, I) = 0 \quad \text{if and only if} \quad U(\mu, I) = 0 .$$

Moreover, by using $(5.2)_\mu$ and the definitions of Y and F one can verify that

$$U(\mu, I) = \operatorname{diag}\left(\frac{1}{\mu(\sigma_i + 1)}\right) Y\left(\mu, \frac{\Lambda}{\mu} - (\sigma + E)I, I\right) = V(\mu, I)I - I ,$$

where

$$V(\mu, I) = \operatorname{diag}\left(\frac{\Lambda_i - (\sigma_i + 1)I_i}{\mu^2(\sigma_i + 1)}\right)\left[\operatorname{diag}\left(\frac{\theta_i\mu}{\Lambda_i - \mu\sigma_i I_i}\right) + \frac{1}{g(\mu, I)} L\right].$$

Lemma 7.3: *$U(\mu, I) \neq 0$, for all $I \in \partial B(\mu)\setminus \{0\}$, $\mu > 0$.*

Proof: Let $(I_1,\cdots,I_n)^T \in \partial B(\mu)\setminus \{0\}$, then one of the following must hold:

(i) There is a j such that $I_j(t) = \dfrac{\Lambda_j}{\mu(\sigma_j+1)}$ and

$$0 \leq I_i(t) \leq \frac{\Lambda_i}{\mu(\sigma_i+1)} \text{ and } \quad 1 \leq i \leq n,\ i \neq j.$$

(ii) $I \neq 0$, $I_i(t) < \dfrac{\Lambda_i}{\mu(\sigma_i+1)}$, $i = 1,\ldots,n$ and there is j such that $I_j = 0$.

Let U_j be the jth component of U and suppose (i) holds. It is then clear that

$$U_j(\mu,I) = -I_j = \frac{\Lambda_j}{\mu\,(\sigma_j+1)} \neq 0,$$

and hence $U(\mu,I) \neq 0$.

Suppose (ii) holds, then $\Lambda_i - \mu\sigma_i I_i \geq \Lambda_i - \mu(\sigma_i+1)\, I_i > 0$, i = 1,..,n, and $g(\mu,I) > 0$. It follows that $V(\mu,I)$ is nonnegative and irreducible since L is nonnegative and irreducible. Thus $U(\mu,I) = V(\mu,I)I - I \neq 0$. Otherwise, in applying [34, Theorem (1.3)], we would have $I \gg 0$, and this contradicts the assumption on I.

Lemma 7.4 ([36, p. 76]): *Let Z be a real Banach space and A be a bounded open set in* $[\alpha, \beta] \times Z$, *and* $\Phi \in C(\overline{A}, Z)$. *Here* $\Phi(\lambda,\mu) = T(\lambda, u) - u$, *and T is continuous. If* $x \in Z \setminus \Phi(\partial A)$ *and* $A_\lambda = A \cap \{\lambda\} \times Z$, *then* $d[\Phi(\lambda, \cdot), A_\lambda, x]$ *is independent of* $\lambda \in [\alpha, \beta]$. *Here* ∂A *refers to the boundary of A in* $[\alpha, \beta] \times Z$ *under the relative topology from* $R \times Z$, *and d denotes the topological degree (for the definition of topological degree, see* [35] *or* [36]).

Definition: *Let*

$$M = \textit{Closure of } \{(\mu, I) \in R_+ \times R_+^n:\ F(\mu, I) = 0\}\ ,$$

$$M^0 = \textit{The maximal connected component of } M \textit{ containing } (\mu_0, 0)\ .$$

M^0 has a clear geometric meaning; it is a global bifurcation branch from the point $(\mu_0, 0)$. Furthermore, by virtue of Lemma 7.2 and 7.3 we see that

$$M^0 \cap \{\mu\} \times \partial B(\mu) = \emptyset \qquad \text{for all } \mu > 0,\ \mu \neq \mu_0\ .$$

Hence

$$M^0 \subseteq \left(\bigcup_{\mu>0} \{\mu\} \times B(\mu) \right) \cup \{(\mu_0, 0)\}\ .$$

We are now ready to establish the following result:

Theorem 7.4: *If* $h(\mu_0) \neq 0$, *then for each* $\mu \in (0, \mu_0)$,

$$M^0 \cap \left(\{\mu\} \times R_+^n\right) \neq \emptyset .$$

Proof: We can prove the theorem by contradiction. Suppose there is a $\xi \in (0, \mu_0)$ such that $M^0 \cap \{\xi\} \times R_+^n = \emptyset$. Since M^0 is connected, by applying Lemma 7.3 we see that

$$M^0 \cap \{\mu\} \times R_+^n = \emptyset, \quad 0 < \mu \leq \xi .$$

Lemma 7.1 implies that M^0 is bounded. Since M^0 and $M \backslash M^0$ are closed, we can find a neighborhood $N \subseteq R_+ \times R^n$ of M^0 such that

$$\overline{N} \cap (M \backslash M^0) = \emptyset, \quad \partial N \cap M = \emptyset . \tag{7.3}$$

Suppose $h(\mu_0) > 0$ [$h(\mu_0) < 0$], then from Theorem 5.2 it follows that there are $\epsilon > 0$ and $0 < \delta_1 < \min\{\Lambda_i/\mu: i = 1,\cdots,n, \mu \in [\mu_0 - \epsilon, \mu_0 + \epsilon]\}$ such that

$$M^0 \cap [\mu_0 - \epsilon, \mu_0 + \epsilon] \times (B_{\delta_1} \cup \{0\}) = \{(\mu, I(\mu): \mu \in [\mu_0 - \epsilon, \mu_0](\mu \in [\mu_0, \mu_0 + \epsilon])\} , \tag{7.4}$$

where $I(\mu)$ is defined as in Theorem 5.2. Now, Lemma 7.2 implies that we can find $\delta > 0$ ($\delta < \delta_1$) such that

$$M \cap [\xi, \mu_0 - \epsilon] \cup \left([\mu_0 + \epsilon, \mu^*] \times (\overline{B_\delta} \backslash \{0\})\right) = \emptyset ,$$

where μ^* is defined in Lemma 7.1.

Insert Figure 1 here

Bifurcation diagram (for details see the text)

(i) Suppose $h(\mu_0) < 0$. Let

$$A_1 = \left(N \setminus \overline{B_1^* \cup B_2^*}\right) \cap \left([\xi, \mu_0 + \epsilon] \times R^n\right),$$

$$A_2 = \left(N \setminus \overline{B_2^*}\right) \cap \left([\mu_0 + \epsilon, \mu^*] \times R^n\right),$$

where

$$B_1^* = [\mu_0 - \epsilon, \mu_0 + \epsilon] \times B_{\delta_1},$$

$$B_2^* = [\xi, \mu_0 - \epsilon] \cup [\mu_0 + \epsilon, \mu^*] \times B_\delta .$$

We can verify that

(a) A_1 is open in $[\xi, \mu_0 + \epsilon] \times R^n$ and A_2 is open in $[\mu_0 + \epsilon, \mu^*] \times R^n$.

(b) Let $A_\mu^i = A_i \cap \{\mu\} \times R^n$, $i = 1, 2$, then by the hypotheses on ξ and (7.3), we have

$$A_\xi^1 \cap \left(M \cup \{(\xi, 0)\}\right) \subset \left(N \setminus \{(\xi, 0)\}\right) \cap M^0 \cap \left(\{\xi\} \times R^n\right) = \emptyset.$$

Similarly,

$$A_{\mu^*}^2 \cap \left(M \cup \{(\mu^*, 0)\}\right) = \emptyset .$$

(c) From the definitions of A_i and N we see that

$$\partial A_\mu^1 \cap \left(M \cup \{(\mu, 0)\}\right) \subset \left(\partial N \setminus \{(\mu, 0)\}\right) \cap M = \phi, \quad \mu \in [\xi, \mu_0 + \epsilon],$$

$$\partial A_\mu^2 \cap \left(M \cup \{(\mu, 0)\}\right) \subset \left(\partial N \setminus \{(\mu, 0)\}\right) \cap M = \phi, \quad \mu \in [\mu_0 + \epsilon, \mu^*] .$$

Now we observe that (c) implies that

$$\partial A_i \cap \left(M \cup [R_+ \times \{0\}]\right) = \emptyset, \quad i = 1, 2 .$$

Thus, by the definition of M we have

$$0 \notin F(\partial A_i) \quad i = 1, 2;$$

or equivalently

$$0 \notin U(\partial A_i).$$

Hence, as a consequence of Lemma 7.4 we have that

$$d\Big(U(u,\cdot),\, A^1_\mu,\, 0\Big) \equiv d\Big(U(\xi,\cdot),\, A^1_\xi,\, 0\Big), \quad \mu \in [\xi,\, \mu_0 + \epsilon]$$

$$d\Big(U(\mu,\cdot),\, A^2_\mu,\, 0\Big) \equiv d\Big(U(\mu^*,\cdot),\, A^2_{\mu^*},\, 0\Big). \quad \mu \in [\mu_0 + \epsilon,\, \mu^*]$$

In addition, (b) implies that

$$d\Big(U(\mu_0 + \epsilon,\cdot),\, A^i_{\mu_0+\epsilon},\, 0\Big) = 0, \quad i = 1, 2\,.$$

On the other hand, we have

$$A^2_{\mu_0+\epsilon} = A^1_{\mu_0+\epsilon} \cup \{\mu_0 + \epsilon\} \times (\overline{B_{\delta_1}} \setminus \overline{B_\delta})\,,$$

and

$$A^1_{\mu_0+\epsilon} \cap \{\mu_0 + \epsilon\} \times (\overline{B_{\delta_1}} \setminus \overline{B_\delta}) = \emptyset\,.$$

By using the properties of degree [36], we deduce that

$$\begin{aligned} 0 &= d\Big(U(\mu_0 + \epsilon,\cdot),\, A^2_{\mu_0+\epsilon},\, 0\Big) \\ &= d\Big(U(\mu_0 + \epsilon,\cdot),\, A^1_{\mu_0+\epsilon},\, 0\Big) + d\Big(U(\mu_0 + \epsilon,\cdot),\, \overline{B_{\delta_1}} \setminus \overline{B_\delta},\, 0\Big) \\ &= d\Big(U(\mu_0 + \epsilon,\cdot),\, \overline{B_{\delta_1}} \setminus \overline{B_\delta},\, 0\Big)\,. \end{aligned} \tag{7.5}$$

Since $\overline{B_{\delta_1}} \setminus \overline{B_\delta}$ contains exactly one zero $I(\mu_0 + \epsilon)$ of $U(\mu_0 + \epsilon,\cdot)$ and

$$\text{sign } det\Big(D_I U[\mu_0 + \epsilon,\, I(\mu_0 + \epsilon)]\Big) = \text{sign } det\Big(D_I F[\mu_0 + \epsilon,\, I(\mu_0 + \epsilon)]\Big) \neq 0\,,$$

hence Lemma 6.2 implies

$$d\Big(U(\mu_0 + \epsilon,\cdot),\, B_{\delta_1} \setminus B_\delta,\, 0\Big) = \text{sign } det\Big(D_I U[\mu_0 + \epsilon,\, I(\mu_0 + \epsilon)]\Big) \neq 0\,,$$

which contradicts the equality (7.5). Therefore, we must have

$$M^0 \cap \Big(\{\mu\} \times R^n\Big) \neq \emptyset, \quad \mu \in (0,\, \mu_0)\,.$$

A similar argument shows that Theorem 7.5 is also true for the case of $h(\mu_0) > 0$.

Corollary 7.6: *If $h(\mu_0) \neq 0$, then for each $\mu \in (0,\, \mu_0)$, $(3.1)_\mu$ has at least one positive equilibrium. Furthermore, if $h(\mu_0) < 0$ and $\epsilon > 0$ are sufficiently small, then for each $\mu \in (\mu_0,\, \mu_0 + \epsilon)$, $(3.1)_\mu$ has at least two positive equilibria.*

Proof: The first conclusion is a direct consequence of Theorem 7.5.

Now suppose $h(\mu_0) < 0$ and $\epsilon > 0$ are defined as in (7.4). In order to prove the second conclusion, it is sufficient to show that $M^0_\mu = M^0 \cap \Big(\{\mu\} \times R^n\Big)$ contains at least two points for each

$\mu \in (\mu_0, \mu_0 + \epsilon)$. First, (7.4) implies that M_μ^0 contains $[\mu, I(\mu)] \in \{\mu\} \times B_\delta$. If our assertion is false for some $\mu \in (\mu_0, \mu_0 + \epsilon)$, then

$$M_\mu^0 \cap B(\mu) \setminus \overline{B_{\delta_1}} = \emptyset . \tag{7.6}$$

Let

$$V_1 = (0, \mu) \times R^n \setminus [\mu_0 - \epsilon, \mu_0 + \epsilon] \times \overline{B_{\delta_1}} ,$$

and

$$V_2 = (\mu_0 - \epsilon, \mu_0 + \epsilon) \times B_{\delta_1} \cup (\mu, \infty) \times R^n .$$

It is clear that V_1 and V_2 are open and disjoint, and (7.6) implies that

$$M^0 \subset V_1 \cup V_2 .$$

Furthermore, (7.4) and Theorem 7.5 guarantee that

$$M^0 \cap V_i \neq \emptyset, \quad i = 1, 2 .$$

This contradicts the connectedness of M^0 and completes the proof.

Remark 1: The first conclusion of the corollary remains true even if $h(\mu_0) = 0$. In fact, it is always possible to alter $h(\mu_0)$ from 0 by making an arbitrarily small change of the other parameters. Then a positive equilibrium can be obtained by taking the limit.

We can also prove this conclusion by using another method. In fact, if we let $T(t)$, $t \geq 0$ be the solution operator of $(3.1)_\mu$, then $T(t)$ is positively invariant and dissipative on $R_+^n \times R_+^n$. Furthermore, we can prove that if $\mu \in (0, \mu_0)$, then $T(t)$ has a maximal, invariant, and compact set $K \subset \Omega = R_+^n \times R_+^n$ which attracts the compact sets of Ω. In applying [32, Lemma 2.6.5, p. 23], one sees that $T(t)$ has a fixed point in K for arbitrary $t > 0$. This implies that $(3.1)_\mu$ has an equilibrium in K.

Remark 2: If $n = 1$ (the case of a single group), then $h(\mu_0)$ is always true. Actually we have proved [21] that whenever $(3.1)_\mu$ has a positive equilibrium; it is unique, and is a global attractor of all positive solutions. But already for the case of more than two groups (i.e. for $n \geq 2$), an example in which $h(\mu_0) < 0$ is not difficult to find. For instance, let $n = 2$, $\mu_0 = 1$, $\sigma_i = 1$, $\Lambda_i = 4$, $\theta_i = 1$,

$r_i = \frac{1}{2}$, i = 1, 2 and

$$L = \begin{bmatrix} 1-\epsilon & 0.1\epsilon \\ 1 & 0.9 \end{bmatrix},$$

where $\epsilon > 0$ is small. Then K = 4 and

$$G(\mu_0, 0) = L - K\,\mathrm{diag}\left(\frac{\mu_0(\sigma_i + 1) - \theta_i}{\Lambda_i}\right) = \begin{bmatrix} -\epsilon & 0.1\epsilon \\ 1 & 0.1 \end{bmatrix}.$$

Let $I_0 = [1, 10]^T$ and $I_0^* = [1/(1 + 10\epsilon), \epsilon/(1 + 10\epsilon)]^T$. It is easy to check that

$$G(\mu_0, 0)I_0 = 0, \quad G^T(\mu_0, 0)I_0^* = 0 \quad \text{and} \quad \langle I_0^*, I_0 \rangle = 1 .$$

Therefore

$$h(\mu_0) = \langle I_0^*, \mathrm{diag}(x_i)I_0 \rangle = \frac{x_1}{1 + 10\epsilon} + \frac{10\epsilon\, x_2}{1 + 10\epsilon} . \tag{7.7}$$

Note that in this case z_1 and x_2 (see Section 5 for the definition of x_i) are independent of ϵ and

$$x_1 = \mu_0\left[\frac{KI_{01}\left(\mu_0(\sigma_1 + 1)^2 - \theta_1\sigma_1\right)}{\Lambda_1^2} - \left(\sum_{j=1}^{2} r_j\sigma_j I_{0j}\right)\frac{\mu_0(\sigma_1 + 1) - \theta_1}{\Lambda_1}\right] = -\frac{2.5}{4} < 0 .$$

It follows from (7.7) that $h(\mu_0) < 0$ if ϵ is sufficiently small. As mentioned in the introduction, if we follow the assumptions made in [10], then we can readily show that $h(\mu_0) > 0$, and hence that the analysis in [10] is not contradicted.

8. Discussion. The system of equations $(3.1)_\mu$ differs from "classical" multigroup epidemic models because the population sizes are not constant, and the nonlinear interaction terms are accordingly more complicated. We have identified the basic reproductive number, which governs the stability of the disease-free equilibrium (an explicit formula in the case of proportionate mixing can be easily be obtained, see [16]), and have obtained an analogue of the usual threshold theorem. What is new here, however, is the possibility of having at least two positive endemic states under some conditions. It is conjectured that one of these is unstable and the other stable. In this case, then, both the disease-free

and one of the endemic states are stable, with different basins of attraction. Whether this situation can occur with realistic parameter values is a matter for further investigation.

Until recently it was believed and supported by numerical simulations that the SIR model had a very simple qualitative behavior. The pressing issues raised by the AIDS epidemic have forced us to re-examine modeling assumptions. The results described in [2], [17-18], and [21], [17] have stressed the importance of epidemiological parameters and variable infectivity. The results of this paper stress the importance of variable population size in disease dynamics. Earlier models for sexually-transmitted diseases (see [1]) assumed that the population and the subpopulations under consideration had a constant number of individuals. Under this assumption the mixing matrix $\{p_{ij}\}$ is time-independent, and consequently a very specific type of mixing is introduced, a mixing that is independent of the population dynamics. Models with varying population size make it possible to study the effects of mixing in population and/or disease dynamics effectively. The results of this paper imply that even in the case of proportionate mixing (see [30] for an example) variable population size can generate multiple endemic equilibria. We have seen that asymmetric mixing patterns are capable of generating multiple endemic equilibria (see Section 7). We suspect that they can occur for realistic parameters. Further studies are required to understand fully the interaction between the mixing matrix and disease dynamics. We and several other investigators have begun to address these and related questions (see [8],[11],[38-40])

Although these results have increased our understanding of the mechanisms behind HIV transmission, they have, unfortunately, revealed a potentially intractable picture. Specifically, if the possible scenarios suggested by the results of this paper and by the results of several other investigators (see [8]) turn out to be generic, then the potential for the use of these models for predictive purposes will be very limited. However, the applicability of these results for theoretical and disease-management purposes looks very promising. Further studies of the effects of mixing patterns in disease dynamics will allow us to better evaluate the role of specific educational programs as well as the possible consequences of possible public policy decisions, including the widespread availability

of prophilactics and the increase availability of drugs such as AZT. Potentially more controversial policy decisions can be *simulated* within the context of mathematical models. We feel, however, that mathematical models cannot be used to circumvent ethical considerations. Finnaly, the issues foreground by this epidemic have led us to new and challenging scientific questions. The models, methods, and techniques that have and will continue to develop will have a significant impact on the growing field of theoretical epidemiology.

References

[1] H.W. Hethcote and J.A. Yorke, *Gonorrhea transmission dynamics and control*, Lecture Notes in Biomathematics 56, Springer-Verlag, Heidelberg, 1984.

[2] R.M. Anderson, G.F. Medley, R.M. May and A.M. Johnson, *A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS,* IMA J. of Mathematics Applied in Med. and Biol., 3, 1986, pp. 229-263.

[3] J.M. Hyman and E.A. Stanley, *Using mathematical models to understand the AIDS epidemic,* Math. Biosci., 90, 1988, pp. 415-473.

[4] R.M. Anderson, *The epidemiology of HIV infection: variable incubation plus infectious periods and heterogeneity in sexual activity,* J. R. Statistic Soc. A, 151, 1988, pp. 66-93.

[5] R.M. May and R.M. Anderson, *The transmission dynamics of human immunodeficiency virus (HIV),* Phil. Trans. R. Soc. London B, 321, 1989, pp. 565-607.

[6] R.M. Anderson, *The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS,* J. AIDS, 1, 1988, pp. 241-256.

[7] S.J. Schwager, C. Castillo-Chavez and H. Hethcote, *Statistical and mathematical approaches in HIV/AIDS modelling: a review,* Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 2-35.

[8] Castillo-Chavez, C. (ed.), *Mathematical and Statistical Approaches to AIDS Epidemiology,* Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989.

[9] J.M. Hyman and E.A. Stanley, *The effect of social mixing patterns on the spread of AIDS,* Mathematical Approaches to Problems in Resource Management and Epidemiology (eds., C. Castillo-Chavez, S.A. Levin and C. Shoemaker), Lecture Notes in Biomathematics No. 81, Springer-Verlag, New York, 1989, pp. 190-219.

[10] J.A. Jacquez, C.P. Simon, J. Koopman, L. Sattenspiel and T. Perry, *Modeling and analyzing HIV transmission: the effects of contact patterns,* Math. Bio., 92, 1988, pp. 119-199.

[11] R.M. Anderson, S.P. Blythe, S. Gupta and E. Konnings, *The transmission dynamics of the*

human immunodeficiency virus type 1 in the male homosexual community in the United Kingdom: the influence of changes in sexual behavior, Phil. Trans. R. Soc. London. B 325, 1989, pp. 45-89.

[12] S. Gupta, R.M. Anderson and R.M. May, *Network of sexual contacts: implications for the pattern of spread of HIV*, AIDS 3, 1989, pp. 1-11.

[13] Blythe, S.P. and C. Castillo-Chavez, *Like-with-like preference and sexual mixing models.* Math. Bio, 96, 1989, pp. 221-238.

[14] C. Castillo-Chavez and S.P. Blythe, *Mixing framework for social/sexual behavior,* Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 275-288.

[15] S. Busenberg and C. Castillo-Chavez, *Interaction, pair formation and force of infection terms in sexually transmitted diseases,* Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 289-300.

[16] S. Busenberg and C. Castillo-Chavez, *On the role of preference in the solution of the mixing problem and its application to risk- and age- structured epidemic models for the spread of AIDS.* Submitted to IMA J. of Mathematics Applied in Med. and Biol.

[17] H.R. Thieme and C. Castillo-Chavez, *On the role of variable infectivity in the dynamics of the human immunodeficiency virus*, Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 157-176.

[18] H.R. Thieme and C. Castillo-Chavez, *On the possible effects of infection-age-dependent infectivity in the dynamics of HIV/AIDS*, 1989, Submitted to J. Of Math. Biol.

[19] L.M. Longini, Jr., W.S. Clark, M. Haber and C.R. Horsburgh, Jr., *The stages of HIV infection: waiting times and infectious contact rates*, Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 111-137.

[20] C. Castillo-Chavez, K. Cooke, W. Huang and S.A. Levin, *On the role of long periods of infectiousness in the dynamics of acquired immunodeficiency syndrome (AIDS),* Mathematical Approaches to Problems in Resource Management and Epidemiology (eds., C. Castillo-Chavez, S.A. Levin and C. Shoemaker), Lecture Notes in Biomathematics, No. 81, Springer Verlag, New York, 1989, pp. 177-189.

[21] C. Castillo-Chavez, K. Cooke, W. Huang and S.A. Levin, *On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS), Part 1. Single population models*, J. Math. Biol., 27(1989), pp. 373-398.

[22] C. Castillo-Chavez, *Recent models of HIV/AIDS transmission*, Applied Mathematical Ecology (eds., S.A. Levin, T.G. Hallam and L.J. Gross), Biomathematics, No. 18, Springer Verlag, New York, 1989, pp. 253-262.

[23] O. Diekmann, J.A.P. Heesterbeek and J.A. Metz, *On the definition of R_0 in models for infectious diseases in heterogeneous populations*, 1989, unpublished manuscript.

[24] D.F. Francis, P.M. Feorino, J.R. Broderson, H.M. McClure, J.P. Getchell, C.R. McGrath, B. Swenson, J.S. McDougal, E.L. Palmer, A.K. Harrison, F. Barre-Sinoussi, J.-C. Chermann, L. Montagnier, J.W. Cuttan, C.D. Cabradilla and V.S. Kalyanaraman, *Infection of chimpanzees with lymphadenopathy-associated virus*, Lancet, 2, 1984, pp. 1276-1277.

[25] J.M.A. Lange, D.A. Paul, H.G. Huisman, F. De Wolf, H. Van den Berg, C.A. Roel, S.A. Danner, J. Van der Noordaa and J. Goudsmit, *Persistent HIV antigenaemia and decline of HIV core antibodies associated with transition to AIDS*, Brit. Med. J., 293, 1987, pp. 1459-1462.

[26] S.Z. Salahuddin, J.E. Groopman, P.D. Markham, M.G. Sarngaharran, R.R. Redfield, M.F. Mclane, M. Essex, A. Gallo and R.C. Gallo, *HTLV-III in symptom-free seronegative persons*, Lancet, 2, 1984, pp. 1418-1420.

[27] S.P. Blythe and R.M. Anderson, *Variable infectiousness in HIV transmission models*, IMA J. of Mathematics Applied in Med. and Biol., 5, 1988, pp. 181-200.

[28] S.P. Blythe and R.M. Anderson, *Distributed incubation and infectious periods in models of transmission dynamics of human immunodeficiency virus (HIV)*, IMA J. of Mathematics Applied

in Med. and Biol., 5, 1988, pp. 1-19.

[29] C. Castillo-Chavez, K. Cooke, W. Huang and S.A. Levin, *On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS), Part 2. Multiple group models*, Mathematical and Statistical Approaches to AIDS Epidemiology (ed., C. Castillo-Chavez), Lecture Notes in Biomathematics, No. 83, Springer Verlag, New York, 1989, pp. 200-217.

[30] C. Castillo-Chavez, K. Cooke, W. Huang and S.A. Levin, *Results on the dynamics for models for the sexual transmission of the human immunodeficiency virus*, Appl. Math. Lett, 2, 1989, pp. 327-331.

[31] A. Nold, *Heterogeneity in disease transmission modelling*, Math. Biosci., 52, 1980, pp. 227-250.

[32] J.K. Hale, *Asymptotic behavior of dissipative systems*, Mathematical Surveys and Monographs, No. 25, Amer. Math. Soc., 1988.

[33] H. Minc, *Nonnegative Matrices*, John Wiley & Sons, New York, 1988.

[34] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[35] S.N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer-Verlag, New York, 1982.

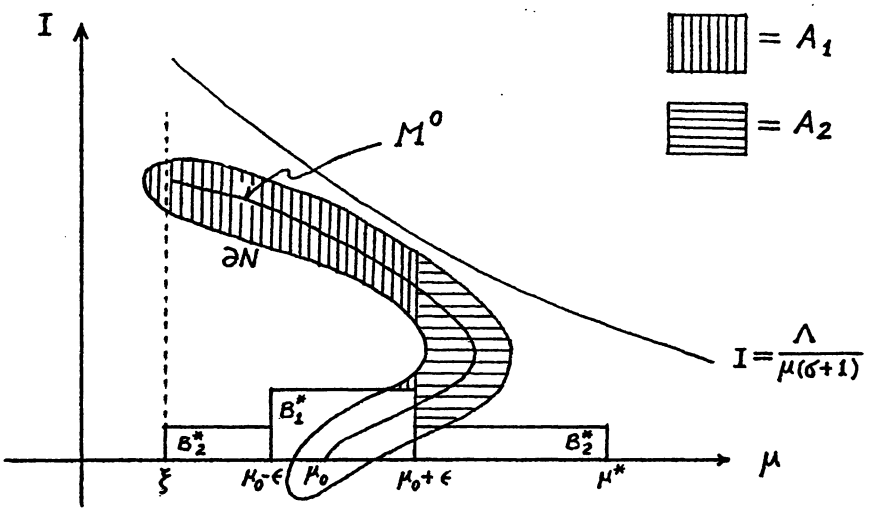
[36] K. Geba and P. Rabinowitz, *Topological Methods in Bifurcation Theory*, Les Presses de L'Universite de Montreal, 1985.

[37] W. Huang, *Studies in differential equations and applications*, Ph.D. dissertation, 1989.

[38] S. P. Blythe, C. Castillo-Chavez and G. Casella, *Empirical methods for the estimation of the mixing probabilities for socially structured populations from a single survey sample*, Submitted to the J. of AIDS.

[39] S. P. Blythe and C. Castillo-Chavez, *On the solution of the pair-formation problem*, unpublished manuscript.

[40] R. M. Anderson, S. Gupta and W. Ng, *The significance of sexual partner contact networks for the transmission dynamics of HIV*, J. of AIDS, 3, 1990, pp. 417-429.


I
$= A_1$
$= A_2$
M^0
∂N
$I = \frac{\Lambda}{\mu(\sigma+1)}$
B_1^*
B_2^*
B_2^*
ξ
$\mu_0 - \epsilon$
μ_0
$\mu_0 + \epsilon$
μ^*
μ